Part II

Features

Architectures for Fault-tolerant Quantum Computing

In theory, quantum computers can efficiently simulate quantum physics, factor large numbers and estimate integrals, thus solving computational problems that are otherwise intractable. In practice, quantum computers must operate with noisy devices called "gates" that tend to destroy the fragile quantum states needed for computation. The goal of fault-tolerant quantum computing is to compute accurately even when gates have a high probability of error each time they are used. We are developing architectural concepts and error remediation strategies which will enable the development of practical quantum computing devices.

Emanuel Knill

Research in quantum computing is motivated by the great increase in computational power offered by quantum computers. There are a large and still growing number of experimental efforts whose ultimate goal is to demonstrate scalable quantum computing. Scalable quantum computing requires that arbitrarily large computations can be efficiently implemented with little error in the output.

One of the criteria necessary for scalable quantum computing is that the level of noise affecting the physical gates is sufficiently low. The type of noise affecting the gates in a given implementation is called the error model. A scheme for scalable quantum computing in the presence of noise is called a fault-tolerant architecture. In view of the low-noise criterion, studies of scalable quantum computing involve constructing fault-tolerant architectures and providing answers to questions such as the following:

> Q1: Is scalable quantum computing possible for error model E?
>
> Q2: Can fault-tolerant architecture A be used for scalable quantum computing with error model E?
>
> Q3: What resources are required to implement quantum computation C using fault-tolerant architecture A with error model E?

To obtain broadly applicable results, fault-tolerant architectures are constructed for generic error models. In such cases, the error model is parameterized by an error probability per gate (or simply error per gate, EPG), where the errors are unbiased and independent. The fundamental theorem of scalable quantum computing is the threshold theorem which answers question Q1 as follows:

> If the EPG is smaller than a threshold, then scalable quantum computing is possible.

Thresholds depend on additional assumptions on the error model and device capabilities. Estimated thresholds vary from below 10^{-6} to 3×10^{-3}, with 10^{-4} often quoted as the target EPG for experimental realizations of quantum computing.

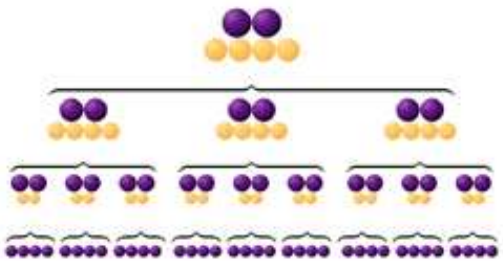

Figure 1. *The new NIST architecture for quantum computing relies on several levels of error checking to ensure the accuracy of quantum bits (qubits). The image above illustrates how qubits are grouped in blocks to form the levels. To implement the architecture with three levels, a series of operations is performed on 36 qubits (bottom row) each one representing a 1, a 0, or both at once. The operations on the nine sets of qubits produce two reliably accurate qubits (top row). The purple spheres represent qubits that are either used in error detection or in actual computations. The yellow spheres are qubits that are measured to detect or correct errors but are not used in final computations.*

Many experimental proposals for quantum computing claim to achieve EPGs below 10^{-4} in theory. However, in the few cases where experiments with two quantum bits (qubits) have been performed, the EPGs currently achieved are much higher, 3×10^{-2} or more in ion traps and liquid-state nuclear magnetic resonance (NMR) experiments, for example.

In our work we have provided evidence that scalable quantum computing is possible at EPGs above 3×10^{-2}. While this is encouraging, the fault-tolerant architecture that achieves this is extremely impractical because of large resource requirements. To reduce the resource requirements, lower EPGs are required.

We have developed a fault-tolerant architecture, called the C_4/C_6 architecture that is well suited to EPGs between 10^{-4} and 10^{-2}. We have analyzed the resource requirements for this architecture and compared it to the state of the art in scalable quantum computing.

The Architecture

Fault-tolerant architectures realize low-error qubits and gates by encoding them with error-correcting codes. A standard technique for amplifying error reduction is concatenation. Suppose we have a scheme that, starting with qubits and gates at one EPG, produces encoded qubits and gates that have a lower EPG.

Provided the error model for encoded gates is sufficiently well behaved, we can then apply the same scheme to the encoded qubits and gates to obtain a next level of encoded qubits and gates with much lower EPGs. Thus, a concatenated fault-tolerant architecture involves a hierarchy of repeatedly encoded qubits and gates. The hierarchy is described in terms of levels of encoding, with the physical qubits and gates being at level 0. The top level is used for implementing quantum computations and its qubits and gates are referred to as being logical. Typically, the EPGs decrease superexponentially with number of levels, provided that the physical EPG is below the threshold for the architecture in question.

The C_4/C_6 architecture differs from previous ones in five significant ways. First, we use the simplest possible error-detecting codes, thus avoiding the complexity of even the smallest error-correcting codes. Error correction is added naturally by concatenation. Second, error correction is performed in one step and combined with logical gates by means of error-correcting teleportation. This minimizes the number of gates contributing to errors before they are corrected. Third, the fault-tolerant architecture is based on a minimal set of operations with only one unitary gate, the controlled-NOT. Although this set does not suffice for universal quantum computing, it is possible to bootstrap other gates. Fourth, verification of the needed ancillary states (logical Bell states) largely avoids the traditional syndrome-based schemes. Instead, we use hierarchical teleportations. Fifth, the highest thresholds are obtained by introducing the model of postselected computing with its own thresholds, which may be higher than those for standard quantum computing. Our fault tolerant implementation of postselected computing has the property that it can be used to prepare states sufficient for (standard) scalable quantum computing.

The properties of the proposed architecture were determined with several months of calculations and simulations on large, conventional computer workstations. Although the new architecture has yet to be validated by mathematical proofs or tested in the laboratory, it provides some evidence that scalable quantum computation may be closer to our reach that previously believed.

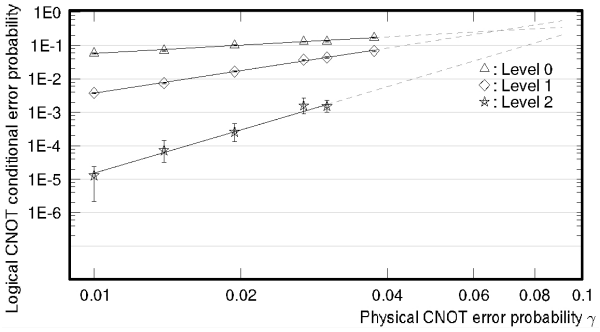

Figure 2. *Errors for a CNOT gate implementation at levels 0, 1 and 2. The errors are conditional on no faults having been detected. The error bars are 68% confidence intervals. As can be seen, errors decrease rapidly with increasing level at EPGs of 3% or below. Extrapolation suggests that this behavior persists for even larger EPGs. At high EPGs, the "no fault" condition happens rarely. Nevertheless it is possible to complete a quantum computation with polynomial overhead by using many trials to prepare relatively error-free states that can then be used to implement error-corrected logical gates with high success probabilities.*

Summary

We have given evidence that accurate quantum computing is possible with error probabilities above 3% per gate, which is significantly higher than what was previously thought possible. However, the resources required for computing at such high error probabilities are excessive. Fortunately, they decrease rapidly with decreasing error probabilities. If we had quantum resources comparable to the considerable resources available in today's digital computers, we could implement non-trivial quantum computations at error probabilities as high as 1% per gate.

References

E. Knill, Quantum Computing with Realistically Noisy Devices, *Nature* **434** (03 March 2005), pp. 39-44.

E. Knill, Scalable Quantum Computing in the Presence of Large Detected-Error Rates, *Physical Review A* **71** (2005), pp. 042322/1-7.

Quantum Logic Circuit Synthesis

Design automation is an important technique for finding efficient classical circuits. Given a Boolean function which the target circuit should implement, a synthesis program automatically determines a sequence of gates realizing this function. Quantum logic synthesis aims to build a similar toolset for quantum circuits, seeking the fewest number of one and two-qubit processes to achieve a target quantum computation. We have developed a technique of automatic quantum circuit synthesis for unstructured qubit evolutions improving a construction from the mid 1990s by a factor of more than one hundred. In addition, we have shown the new construction to be within a factor of two of optimal. For multi-level quantum logics (qudits) similar advances have led to the first circuits ever with optimal asymptotics.

Stephen S. Bullock

While classical computers manipulate bits which carry values of 0 or 1, quantum computers manipulate quantum bits (qubits) which are state vectors of two-level quantum systems. If the quantum computer is not exchanging energy with the outside environment, these qubit state vectors are rotated during the computation. Typical quantum algorithms call for implementing such a rotation, mathematically a unitary matrix, and then observing the qubits. Thus, while efficient Boolean circuits realize complicated functions on bits using a small number of logic gates, efficient quantum circuits break complicated unitary matrices into simple factors. These factors (quantum gates) typically correspond to manipulating one or two quantum bits.

The past two years have seen marked advances in the design of universal quantum logic circuits. Such circuits implement any possible unitary evolution by appropriately tuning their gate parameters. The new techniques are also overtly constructive. Unitary matrices implementing nontrivial quantum computations are large, e.g. requiring $2^n \times 2^n$ matrices for n qubits. The new quantum circuit synthesis algorithms rely on well-known matrix decomposition such as QR or the Cosine Sine Decomposition. For ten qubits, commercial software on a 2.5GHz PC requires a few seconds for these factorizations.

We illustrate basic quantum circuit design with an example. Each circuit in the figure below applies the same two-qubit computation, namely multiplying the *01* and *10* states by the complex number i while leaving *00* and *11* unchanged. Each qubit is represented by a single line or rail in the circuit. The boxes denoted S and H indicate single qubit operations, i.e. particular 2×2 unitary matrices. The gate spanning both qubits is a quantum controlled not, or CNOT. CNOT flips the target qubit (carrying the inverter) when the control qubit (black slug) carries *1*, so that these down-target CNOTs exchange *10* and *11*.

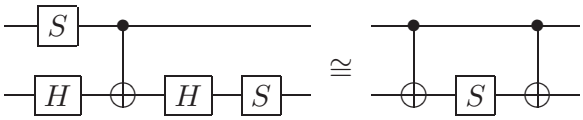

Figure 3. *Two quantum circuits for performing the same task.*

Although their target computation is the same, either diagram might be better suited to a given quantum computer implementation. For example, if CNOTs are costly while one-qubit gates are implemented more easily, then the circuit at left is preferable. Alternately, if one-qubit gates are expensive but CNOTs are cheap, then the circuit at right is preferable. In practice, CNOTs tend to be more expensive than one-qubit rotations, and three-qubit gates tend to be more difficult to implement than CNOTs. Further, while it is possible to build any unitary using exclusively one qubit gates and CNOT, simply employing one-qubit gates does not suffice. Thus, the following discussion of universal quantum circuits focuses on minimizing CNOT counts.

In fact, it is not obvious that CNOT and one-qubit gates suffice to build any unitary evolution. This was settled in 1995 in a landmark paper of Barenco, Bennett, Cleve, DiVincenzo, Margolus, and Shor. If n is the number of qubits on which the unitary operator acts, then these authors showed that $48n^3 4^n$ CNOTs suffice, in addition to many one-qubit gates. Shortly thereafter, Knill argued that some multiple $C4^n$ gates must be required for reasons of dimension. The result has a parallel in classical circuits. Namely, given a *random* bit-valued function on n bit strings, approximately $2^n/2$ strings will take on a value of 1. Thus, we expect $2^n/2$ gates are required to distinguish for which bit strings the circuit should return a nonzero value.

A Unitary-Universal *n* Qubit Circuit

In summer of 2004, we discovered a new unitary-universal n qubit quantum circuit requiring roughly $(1/2)4^n$ CNOTs, an improvement by a factor of two over the best known circuit at the time and a factor of $100n^3$ over the 1995 circuit. Moreover, work from the summer of 2003 had sharpened Knill's bound, actually valid for any two-qubit gate, to a specific bound of $(1/4)4^n$ CNOTs. Hence, the present circuit may never be improved by more than a factor of two.

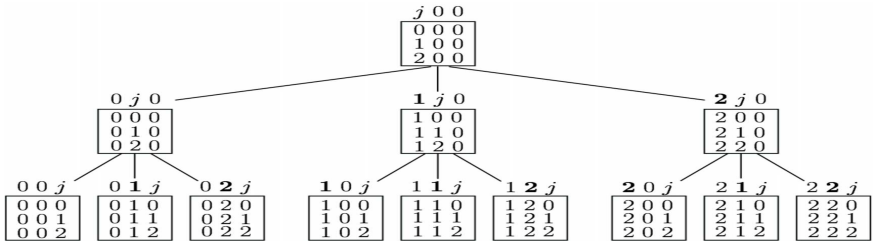

Figure 4. *Recursively generated tree for automated quantum circuit synthesis for three qutrits.*

The outline for deriving the circuit is as follows. The key step is to use the Cosine-Sine Decomposition (CSD) for matrices. The CSD splits any unitary matrix into three factors; the circuit elements outline in the box below represent the first and last. The slash in the circuit represents a multi-line carrying an arbitrary number of qubits, meaning U may be any unitary matrix on any number of qubits. Moreover, the circuit elements having the square box controls correspond to uniformly controlled rotations, a circuit block for which particularly CNOT-efficient circuits are known. Hence, the diagram employs the CSD to reduce construction of the n qubit U into four simpler n-1 qubit unitaries V_1, V_2, V_3, and V_4. This allows for a recursive construction which terminates with hand-optimized two-qubit circuits.

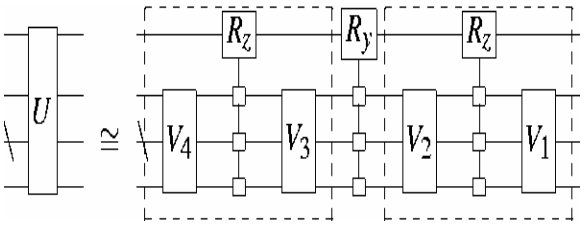

Figure 5. Basic quantum circuit synthesis decomposition.

Circuits for Quantum Multi-Level Logics

In addition, the efficiency of the circuit for quantum multi-level logics (qudits, d=level) has improved. One may think of a qudit by way of analogy: a qubit state is a quantum superposition of *0* and *1*, a qudit state is a quantum superposition of *0, 1, … d-1*. Knill's lower bound also stated that Cd^{n-2} two-qudit gates were required for any generic $d^n \times d^n$ unitary U. Yet the best known constructive procedure required Cn^2d^{2n} gates. In fall of 2004, we produced a new Cd^{2n} construction, closing this gap. The circuit exploits a variant of the QR matrix decomposition. The original asymptotically optimal qubit circuits, due to researchers at the University of Helsinki, also leaned on a QR technique and a Gray code cancellation. Rather than generalize the Gray code cancellation to base d numbers, the new circuit relies on a recursively generated tree. This tree describes which entries of the unitary matrix to construct with quantum gate elements at which time. A sample tree, for 3 qutrits (i.e., d=*3*, n=*3*) is shown above.

Lessons Learned

These specialized techniques for generic unitary matrices might help in optimizing other quantum circuits.

- Just as Boolean factorizations are important for classical circuit design, so too are matrix decompositions useful in quantum circuit design.
- It is possible to exploit parallels to classical logic synthesis. For example, one may view the side factors of the CSD as multiplexers, applying an *n-1* qubit unitary matrix dependent on the most significant qubit.
- Novel aspects of quantum circuits must be explored thoroughly. For example, at the heart of the qudit circuit is a subcircuit capable of solving the hard problem of initializing a generic quantum memory state. In contrast, the classical problem of initialization trivially requires at most n bit flips.

References

V.V. Shende, S.S. Bullock, I.L. Markov, A Practical Top-down Approach to Quantum Circuit Synthesis, Proc. *Asia and South Pacific Design Automation Conference*, pp. 272-275, Shanghai, China, 2005.

S. Bullock D.P. O'Leary and G. Brennen, Asymptotically Optimal Quantum Circuits for d-level Systems, *Physical Review Letters* **94**, 230502 (17 June 2005).

Participants

S. Bullock, D. P. O'Leary (MCSD), G. Brennen (PL), V. Shende, I. Markov (Univ. of Michigan)

http://math.nist.gov/quantum/

Adaptive Finite Element Modeling of Two Confined and Interacting Atoms

High order finite element methods using adaptive refinement and multigrid techniques have been shown to be very efficient for solving partial differential equations on sequential computers. We have developed a code, PHAML, to extend these methods to parallel computers. We have applied this code to solve a two-dimensional Schrödinger equation in order to study the feasibility of a quantum computer based on extremely cold neutral alkali-metal atoms. Qubits are implemented as motional states of an atom trapped in a single well of an optical lattice. Quantum gates are constructed by bringing two atoms together in a single well leaving the interaction between them to cause entanglement. Quantifying the entanglement reduces to solving for selected eigenfunctions of a Schrödinger equation that contains a Laplacian, a trapping potential, and a short-range interaction potential.

William Mitchell

The idea of using the rules of quantum mechanics as a paradigm for computing has engendered a flurry of research over the last ten years. Strange quantum properties, such as entanglement, may yield a significant advantage, providing novel mechanisms for the solution of problems that are intractable on classical computers. Advances in diverse fields of physics have led to proposals for various alternate physical realizations of a quantum bit and related quantum gates, the quantum analog of one- and two-bit computer operations. We are interested in modeling a quantum gate with quantum bits that are based on ultra-cold atoms. Ultra-cold atoms can be confined by counter-propagating laser beams. The light creates a three-dimensional washboard potential or optical lattice. A single atom is held in each potential minimum or lattice site of the washboard. Two energy levels of an atom are associated with the "0" and "1" states of a quantum bit. By bringing two isolated atoms from separate sites together and having them interact, a two-quantum-bit operation can be realized.

First-principle modeling of the interactions of two atoms in a lattice site is numerically challenging, as length scales for the lattice and the mutual atom-atom interaction differ by orders of magnitude. Our work has applied advanced finite element techniques, including high order elements, adaptive grid refinement, multigrid solution methods, and parallel computing, to the solution of the Schrödinger equation that models this interaction.

For appropriately chosen laser intensities and geometries the lattice sites are approximately harmonic and cylindrically symmetric. This leads to a model for the relative motion of two atoms in a single site by a two-dimensional Schrödinger equation in cylindrical coordinates, i.e., an elliptic eigenvalue problem. The potential function in the Schrödinger equation consists of a short-ranged atom-atom interaction potential, modeled by a Lennard-Jones potential, and a trapping potential that describes the optical well.

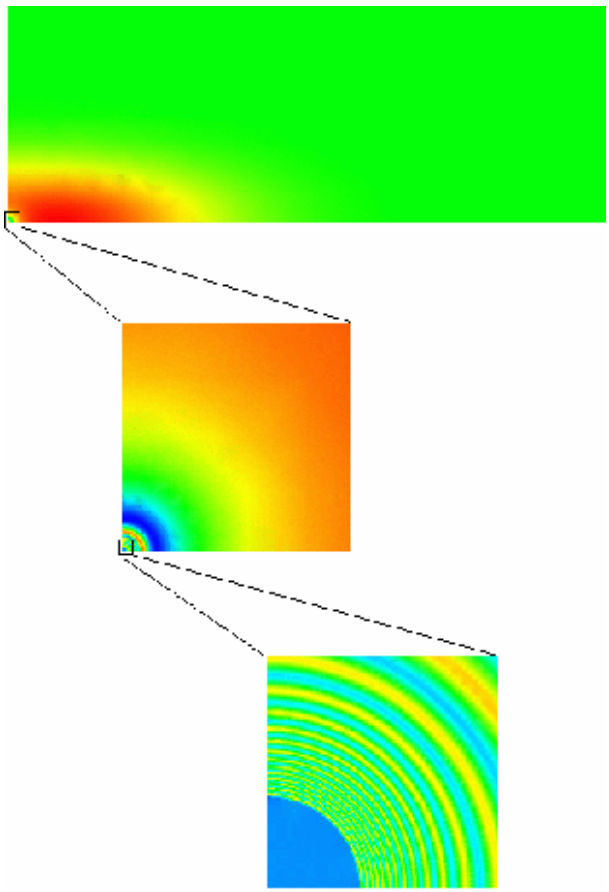

Figure 6. *Contour map of the computed wave function for the first trapping state of a model of interacting Cesium atoms. The large difference in scales between different parts of the wave function is illustrated by two levels of zooming.*

The interest lies in obtaining a small number of wave functions whose eigenvalues are closest to zero. These are called the trap states because they are eigenstates in which the eigenfunction extends from short to large atomic separations and in which the trapping potential plays an important role. Fig. 6 illustrates the nature of

the wave function for the first trapping state of a model of Cesium atoms.

The large difference between the scales of the wave due to the trapping potential, the large red wave in Fig. 6, and the waves due to the interaction potential, the small waves, which are evident through two levels of zooming the image, mandate the use of adaptive grid techniques. An adaptive grid technique begins with a very coarse grid, and then selectively refines elements by computing an error estimate for each element and refining those with large estimates. The process is repeated until a sufficiently accurate solution can be computed. The grid used for the solution shown in Fig. 6 consisted of approximately three million elements. The elements are extremely small in the area where the solution varies rapidly, and large in the outer regions where the solution varies slowly. Such an element distribution can lead to optimal utilization of resources during the solution process. If a uniform grid with elements the size of the smallest elements was used, for example, it would require on the order of 10^{17} elements, clearly showing the need for non-uniform grids for this problem.

Additional improvements in the solution have recently been made through the application of high order elements. Linear elements represent the solution as a piecewise smooth function that is linear over each triangle of the grid. High order elements use a higher degree polynomial on each triangle. Using a p^{th} degree polynomial, the error of the approximation decreases like h^{p+1} where h is the diameter of the element. Thus high order elements give an accurate solution with many fewer grid elements than in the linear case. For another Cesium model, a` solution with 3 accurate digits via linear elements required approximately 4.5 million degrees of freedom. The solution took 35 minutes on a 32-processor parallel computer. Using cubic elements, the same accuracy is obtained with 0.5 million degrees of freedom in 8 minutes on a laptop computer. However, we require much higher accuracy for realistic models. Using 5^{th} degree elements and resources comparable to the linear case, we have obtained solutions accurate to 8 digits.

Further improvements in the solution techniques will be made through our on-going research in the use of hp-adaptivity. In this approach, adaptive refinement is applied not only to the size of elements (h), but also to the polynomial degree over each element (p). The appeal of hp-adaptivity is that the error of the approximation can decrease exponentially in the number of degrees of freedom, whereas with fixed degree polynomials it can only decrease polynomially. Fig. 7 shows the exponential rate of convergence we have obtained for a model elliptic boundary value problem using hp-adaptivity. Polynomial convergence would appear as a straight line on this graph. The curvature of the line obtained by a least squares fit to the data illustrates that exponential convergence has been obtained. In our current research we are developing new error estimates that will allow us to apply hp-adaptive techniques to elliptic eigenvalue problems.

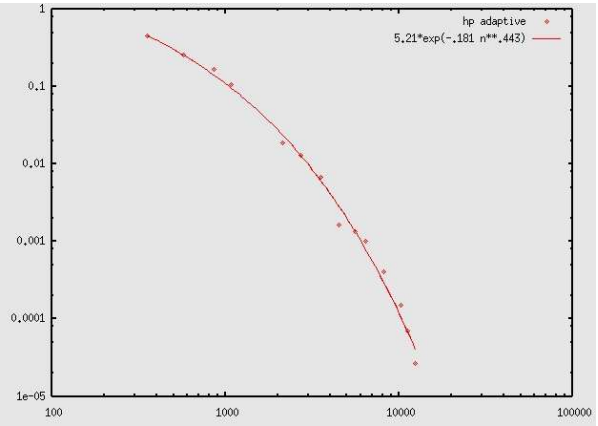

Figure 7. *Exponential convergence of the error vs. degrees of freedom for the hp-adaptive solution of an elliptic boundary value problem.*

With the recent addition of high order elements to our models, we are now beginning to perform numerical experiments on systems relevant to realistic atom traps. Currently we are investigation the effect of varying the scattering length of the interaction potential and eccentricity of the trapping potential on a model of Cesium atoms. We anticipate that the addition of hp-adaptivity will further reduce the solution time to the point where we can perform experiments with the multi-channel time-dependent equations that are required for realistic models of quantum gates.

References

W.F. Mitchell and E. Tiesinga, Adaptive Grid Refinement for a Model of Two Confined and Interacting Atoms, *Applied Numerical Mathematics* **52** (2005), pp. 235-250.

W.F. Mitchell, Error Estimators for the hp Version of the Finite Element Method with Newest Node Bisection of Triangles, 8th U.S. National Congress on Computational Mechanics (July 2005).

Participants

W. Mitchell (MCSD); E. Tiesinga (PL)

Mathematical Modeling of Nanomagnetism

Measurement, understanding, and control of magnetic phenomena at the nanoscale each require the support of mathematical models of the physics involved, and software that correctly implements and makes predictions based on these models. The Object-Oriented MicroMagnetic Framework (OOMMF) project provides this capability in a public domain package of portable software components organized in an extensible framework. OOMMF software is widely used and cited in the physics and engineering literature. Current objectives are to continue to expand the features supported by components in the OOMMF framework, with focus on high priority items, such as thermal effects and spin transfer, motivated by their relevance to nanoscale sensor and spintronics work.

Michael J. Donahue and Donald G. Porter

Many existing and developing applications of nanotechnology make use of magnetic phenomena. Some of the most familiar and successful examples are information storage technologies such as magnetic recording media, GMR sensors for read heads, and magnetic RAM (MRAM) elements. Computational modeling continues to support advances such as the novel patterned magnetic recording media that promise to achieve recording densities of 1 Tb per square inch.

Applications of nanomagnetism modeling are widespread. For example, modeling was critical to the development of a fully magnetic logic gate and shift register accomplished at University of Durham and recognized by the Institute of Physics as one of the top ten stories in physics for 2002. Other efforts aim to develop logic devices that carry information in the form of the spin of a charge carrier, so called *spintronics* devices. In biotechnology, the use of paramagnetic beads to locate and position biological macromolecules is under study. In materials science, the probing capability of ferromagnetic resonance is used to characterize and measure material properties, where an understanding of nanomagnetodynamics is a key to interpreting experimental results. Other efforts in materials science are aimed at improved sensor designs capable of detecting magnetic fields that are both smaller in magnitude and more localized in space.

In each of these areas, the NIST Object-Oriented MicroMagnetic Framework (OOMMF) system is in use to enable nanoscale science and engineering. OOMMF is a portable public domain package organized in an extensible framework to enable computational simulation of magnetic systems.

One recent example of the use of OOMMF to support nanotechnology R&D is depicted in Fig. 8. An image of 2000 nm diameter ring of ferromagnetic material was produced by a magnetic force microscope (MFM). Such ring elements are proposed for both storage and sensor applications, where their utility is critically dependent on their precise behavior. The colors of the image represent the strength of stray magnetic field sensed at each location scanned above the ring. An MFM directly measures stray field; it does not directly measure the magnetization pattern found in the ring itself. A measured stray field does not uniquely determine what magnetization pattern produced the stray field. Also illustrated is a magnetization pattern computed by OOMMF software based on known parameters chosen to match the experimental work. The pattern of arrows represents the magnetization pattern predicted by the model, and it can be confirmed that the predicted pattern is consistent with the measured stray field. In complementary roles like this, modeling is able to indirectly deduce details of a nanoscale system that are not within the capability of direct measurement.

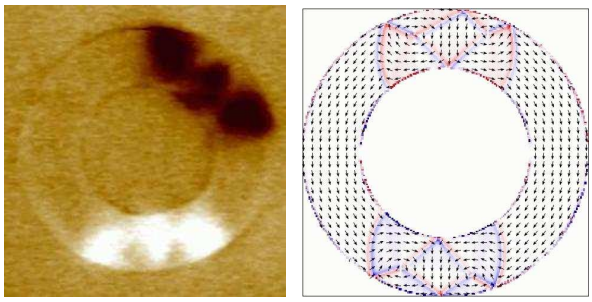

Figure 8. *Left: Image of 2000 nm diameter ring of ferromagnetic material produced by a magnetic force microscope. Right: Magnetization pattern computed by OOMMF based on known parameters chosen to match the experimental work.*

Many of the equations governing nanomagnetism were established long ago. The notable Landau-Lifshitz equation dates back to 1935, and William Fuller Brown established the fundamentals of *micromagnetic modeling* in articles and books published from the 1940s to 1970s. At that time, the applications of the theory were mostly limited to simple geometries that could be attacked analytically. It was not until the 1990s that widespread availability of significant inexpensive computing power made possible micromagnetic modeling for practical problems.

Unfortunately, the accurate solution of the relevant equations is more difficult than many physicists and

engineers realize. By the mid-1990s, many researchers and their students had produced their own simulation programs to support their work. While their publications were careful to list details of experiments and analyses, their home-grown software would typically only merit a high-level description. Usually the equations the software was meant to solve would be noted, but no reports of software testing were provided, and there was no opportunity to review its quality. The problems solved by these computer codes were typically precisely matched to a particular experiment, and it was rare that two independent teams would perform precisely the same computations to enable comparison.

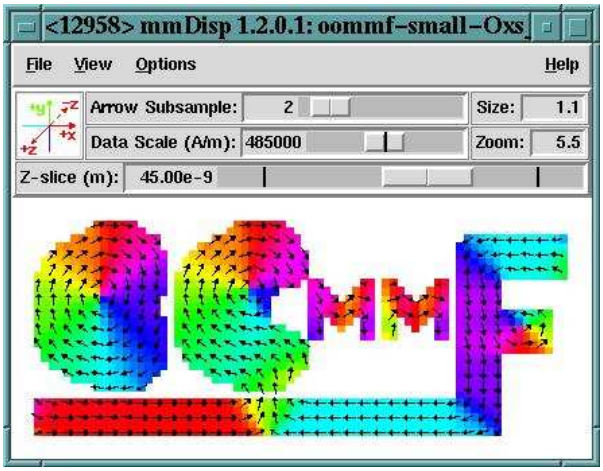

Figure 9. *Part of the graphical user interface for OOMMF.*

In 1996, NIST challenged this research community with a *standard problem*, inviting all those with magnetic modeling simulation software to attempt to compute some properties of a magnetization reversal similar to those routinely reported in the literature at the time, but all starting from the same assumptions. The results were alarming. The number of significantly different solutions almost matched the number of programs. Subsequent research has discovered flaws in both the programs and the problem. An important value of OOMMF is that it provides a transparent benchmark against which any research team developing their own magnetic modeling software can compare their results.

Those researchers for whom OOMMF provides all required modeling capabilities can, of course, also use OOMMF in place of developing their own software. In addition, OOMMF is structured as an extensible framework of software components, so those users who have needs beyond OOMMF's current capabilities can often extend it to meet their needs without the need to rewrite a new software package from scratch.

OOMMF has been remarkably successful in achieving its objectives. OOMMF has been downloaded more than 10,000 times. A growing number of peer-reviewed research publications cite use of OOMMF; more than 300 are listed on the OOMMF web site.

Current plans for OOMMF are to supply additional extension modules and to make improvements to the framework necessary to support the nanomagnetic simulations required in emerging research areas. For example, NIST research in the development of ultra-low field magnetic sensors requires reliable modeling of thermal effects, a component that OOMMF has until now not provided. There is also considerable interest in representing the effects of spin transfer in nanomagnetic systems. In each of these cases, research is active into determining the correct models.

References

N. Dao, M.J. Donahue, I. Dumitru, L. Spinu, S.L. Whittenburg, and J.C. Lodder, Dynamic Susceptibility of Nanopillars, *Nanotechnology* **15** (2004), pp. S634-S638.

L. Yanik, E. Della Torre, M.J. Donahue, and E. Cardelli, Micromagnetic Eddy Currents in Conducting Cylinders, *Journal of Applied Physics* **97** (2005), p. 10E308.

E. Mirowski, J. Moreland, A. Zhang, S.E. Russek, and M.J. Donahue, Manipulation and Sorting of Magnetic Particles by a Magnetic Force Microscope on a Microfluidic Magnetic Trap Platform, *Applied Physics Letters* **86** (2005), 243901.

M.H. Park, Y.K. Hong, B.C. Choi, S.H. Gee, and M.J. Donahue, Vortex Head-to-head Domain Walls and Its Formation Process in Onion-State-Stable Ring Elements, submitted.

Participants

M. Donahue, D. Porter (MCSD); R. McMichael (MSEL)

`http://math.nist.gov/oommf/`

Stability of Nanowires

Metallic or semiconductor nanowires are important components in many electronic technologies. At small length scales many materials exhibit unusual electrical, chemical, and thermal properties that are not observed in the bulk. Applications include novel electronic devices ranging from high efficiency lasers and detectors to exotic single electron transistors and cellular automata. The tendency of nanowires to fragment into nanospheres due to area-minimizing surface instabilities acts as a limit to the length of nanowires that can be used in nanodevices. This instability can also be used beneficially as a mechanism for the self-organization of chains of nanospheres from unstable nanowires. We have studied analytically the effect of surface tension anisotropy on this instability in order to help understand and control nanowire fragmentation. Our analysis predicts the wavelengths of the instability as a function of the degree of anisotropy of the surface energy of the nanowire.

Geoffrey McFadden

At the small length scales that characterize nanostructures, the importance of surface effects relative to volume effects becomes significant. Typical surface effects that can be important at the nanoscale include surface energy or capillarity, surface diffusion, surface adsorption, and surface stress and strain. In particular, the effects of capillarity must be taken into account in order to understand the tendency of nanowires to fragment when the rate of surface diffusion of atoms is high enough to allow shape changes to occur over practical time scales.

As shown by Plateau in his classical studies of capillary instabilities, a cylindrical interface with an isotropic surface free energy is unstable to volume-preserving axisymmetric perturbations whose wavelength exceeds the circumference of the cylinder. Such perturbations lower the total energy of the cylinder, leading to the breakup of the cylinder into a series of drops or bubbles. The stability of a liquid jet was subsequently studied by Lord Rayleigh, who argued that the length scale of the instability is determined by the perturbations having the fastest growth rate; the phenomenon has generally come to be known as the Rayleigh instability. The Rayleigh instability arises in a number of diverse applications, such as ink jet printing, two-phase flow, quantum wires, fiber spinning, liquid crystals, and polymer blends.

Because of the underlying crystal lattice, the surface energy of a liquid-solid or vapor-solid interface is generally anisotropic and depends on the orientation of the local normal vector at each point of the interface. The surface energy of a solid-solid interface between two crystals is also anisotropic in general, with the additional complication that the surface energy also depends on the direction cosines that characterize the relative orientations of the two crystals. Here we have considered a model of this type in which the surface energy depends only on the local normal vector.

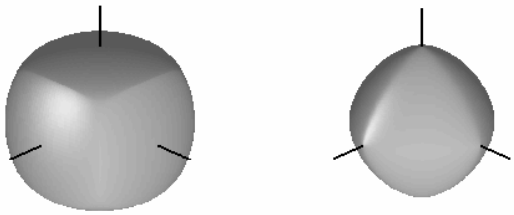

Figure 10. *Equilibrium shapes for materials with cubic anisotropy. These shapes are energy minimizing surfaces that are the anisotropic versions of soap bubbles (the isotropic case).*

An observation that partially motivates this work is the apparent stability of elongated nanowires that are grown in a bridge configuration or epitaxially on a heterogeneous substrate. The nanowires (alternatively called nanorods or quantum wires) are "one-dimensional" crystals with dimensions as small as one nanometer high, a few nanometers wide, and can be as long as a micron. There are long-standing studies on experimental techniques to grow nanowires, and the stability of these nanowires is beginning to come under study. Another motivation for the work is the recent observation that the Rayleigh instability of a nanowire can be used to produce self-organized chains of nanospheres with interesting electrical and opto-electronic properties. In either case it is desirable to develop models to predict the length scales of the instabilities in order to assess the geometry of the resulting structures.

Continuum modeling of nanowires provides some guidance as to their expected stability, though the strict applicability of continuum models is limited if the length scales approach atomic dimensions. There are a number of possible mechanisms that could stabilize a nanowire, including elastic interactions between the wire and the substrate, quantum electronic shell effects, and surface energy anisotropy. A useful model of the surface energy anisotropy for a cubic material is given by the expression $\gamma(n_x,n_y,n_z) = \gamma_0 [1 + 4 \varepsilon (n_x^4 + n_y^4 + n_z^4)]$. In the above figure we show examples of 3D equilibrium shapes co responding to this surface energy. The shapes are smooth for $-1/18 < \varepsilon < 1/12$. For $\varepsilon < 0$, the shapes resemble rounded cubes, with

[110] edges first forming at $\varepsilon = -1/18$. As ε decreases below $-1/18$, the edges extend toward the [111] directions, merging to form a corner for $\varepsilon = -5/68$. For $\varepsilon > 0$ the shapes are octahedral, with [100] corners first forming at $\varepsilon = 1/12$. These equilibrium shapes are most easily computed using the ξ-vector formalism of Hoffman and Cahn, which produces a closed-form expression for the equilibrium shape in terms of the surface energy. This is also useful in formulating the variational problem for the stability of a nanowire.

In collaboration with K. Gurski, George Washington University, and M. Miksis, Northwestern University, we have considered differentiable surface energies with anisotropies mild enough that the surface of the wire is smooth and does not exhibit any missing orientations. In order to examine the stability of the wire using a variational approach, we employ a general energy functional that describes the total surface energy of the system. This expression and the constraint of constant volume of the wire are perturbed about the two-dimensional equilibrium shape. The higher order terms in this perturbation expansion produce a condition for stability. For constant volume, if the perturbation increases the energy, the equilibrium state is stable, otherwise it is unstable.

Figure 11. *Capillary instability of a nanowire with an anisotropic surface energy with three-fold symmetry about the wire axis.*

For small levels of anisotropy, we evaluate the stability of an isolated nanowire approximately using asymptotics. For larger amplitudes of anisotropy, we compute solutions numerically. We find that surface tension anisotropy can either promote or suppress the Rayleigh instability, depending on the orientation of the nanowire and the magnitude and sign of the anisotropy. For general surface energies we derive an associated eigenproblem whose eigenvalues govern the stability of the wire. The eigenproblem is described by a pair of coupled second-order ordinary differential equations with periodic coefficients, which generally lack closed-form solutions. We have applied the analysis to a number of examples, including the above case of a cubic material. We have computed the stability of the wire to general perturbations when the axis of the wire is in a high symmetry orientation such as [001], [011], or [111].

In addition to determining the stability of an isolated wire, we have also examined how both the anisotropy of the surface energy of the wire and the interaction of the wire with a substrate affects the stability of the rod. The equilibrium configuration of a wire in contact with a substrate has an elegant description that can be obtained by again appealing to the Hoffman-Cahn ξ-vector formalism. This approach determines the contact angles in terms of the surface energies of the phases that meet at the contact line. Using general anisotropic surface energies we have then derived an associated eigenproblem that describes the stability of the system. The problem is described by a pair of coupled second-order ordinary differential equations with periodic boundary conditions along the axis of the rod and boundary conditions arising from the contact angles between the wire and substrate. We have considered the effects of the overall orientation of the crystal relative to the substrate and examined a range of contact angles. The substrate is assumed to be rigid with an isotropic surface energy.

We applied the analysis to a number of examples, including the case of a cubic material, and compute the stability of the wire to perturbations when the axis of the wire is aligned parallel to the high symmetry orientations [001], [011], and [111]. We assumed a weak anisotropic surface energy to eliminate missing orientations on the wire. The magnitude and the sign of the anisotropy determine the relative stability in comparison to the isotropic case. In general as the contact angle tends to 90 degrees the wire becomes more stable, which is analogous to the stability of a 3D planar film.

References

K.F. Gurski and G.B. McFadden, The Effect of Anisotropic Surface Energy on the Rayleigh Instability, *Proceedings of the Royal Society (London) A* **459** (2003), pp. 2575-2598.

K.F. Gurski, G.B. McFadden, and M.J. Miksis, The Effect of Contact Lines on the Rayleigh Instability with Anisotropic Surface Energy, *SIAM Journal on Applied Mathematics,* in press (2005).

Participants

G. McFadden (MCSD); K. Gurski (George Washington Univ.); M. Miksis (Northwestern Univ.)

Improving Image Resolution in Nanotechnology

Current nanoscale electron microscopy images remain of relatively low quality. To address this issue we are developing improved mathematical tools for image analysis, and the use of such tools to provide measurable increases in resolution in state-of-the-art scanning electron microscopy. One very major difficulty lies in the large image sizes, often on the order of 1024×1024 pixels, or larger. This presents formidable computational challenges. Many new techniques are based on nonlinear partial differential equations, and typically require thousands of iterations, and several hours of CPU time, to achieve useful results. Real-time image processing algorithms are exceedingly rare and very highly sought after.

Alfred Carasso

A fundamental problem in scanning electron microscopy (SEM) is the fact that the shape of the electron beam that produced the image is seldom known to the microscopist. Therefore, image deblurring must proceed without knowledge of the actual point spread function that caused the blur. Such so-called *blind deconvolution* is fraught with difficulty, and little authoritative discussion of this subject is to be found in most image processing textbooks.

Nevertheless, in recent years, considerable progress was achieved at NIST in developing mathematical technologies that lead to real-time image processing algorithms. In addition, a unique new capability has been created, the so-called *APEX method*, that can achieve useful blind deconvolution of 1024×1024 SEM imagery in about 60 seconds on current workstations. Because of its manifold applications, this technology is the subject of intense and continuing research and development.

The APEX Method

The APEX method is an FFT-based direct blind deconvolution technique that can process complex high resolution imagery in seconds or minutes on current desktop platforms. The method is predicated on a restricted class of shift-invariant blurs that can be expressed as finite convolution products of two-dimensional radially symmetric Lévy stable probability density functions. This class generalizes Gaussian and Lorentzian densities but excludes defocus and motion blurs. Not all images can be enhanced with the APEX method. However, we have shown that the method can be usefully applied to a wide variety of real blurred images, including astronomical, Landsat, and aerial images, MRI and PET brain scans, and SEM images. APEX processing of these images enhances contrast and sharpens structural detail, leading to noticeable improvements in visual quality.

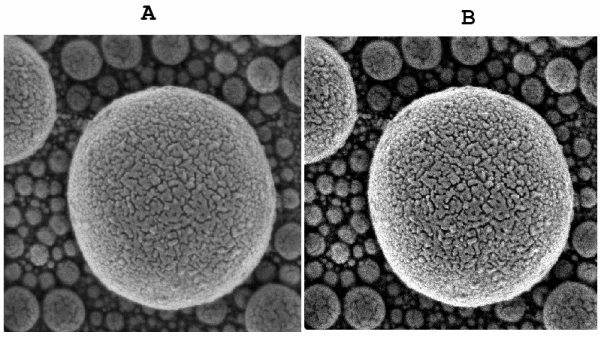

Figure 12. *APEX blind deconvolution of state of the art Scanning Electron Microscope imagery produces measurable increases in sharpness. (A) Original 1024x1024 Tin sample micrograph has Lipschitz exponent* $\alpha = 0.40$ *and TV norm = 13000. (B) Sharpened image has* $\alpha = 0.29$ *and TV norm = 34000.*

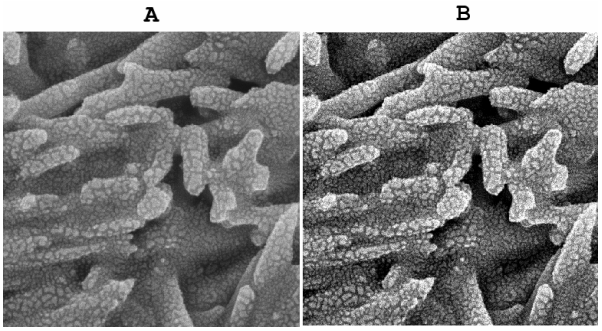

Figure 13. *APEX sharpening of SEM imagery. (A) Original 1024x1024 Magnetic Tape sample has* $\alpha = 0.35$ *and TV norm = 14000. (B) Sharpened image has* $\alpha = 0.26$ *and TV norm = 39000*

Application to SEM Imagery

Recently, a new Hitachi Scanning Electron Microscope was acquired by the NIST Nanoscale Metrology Group, capable of producing higher quality imagery than had previously been possible. A major challenge for our deconvolution algorithms was to demonstrate measurable increases in sharpening of such state of the art imagery. Two sharpness measures were used, the image Lipschitz exponent α, and the image discrete *total variation* or TV norm. Image sharpening increases the TV norm, due to the steepening of gradients, while it decreases the Lipschitz exponent as finer scale features become resolved. Examples of such sharpening are shown in Figs. 12 and 13. In Fig. 12A, the original *1024×1024* Tin sample micrograph has TV norm of 13000 and Lipschitz exponent $\alpha = 0.40$. The APEX-sharpened Fig. 12B has TV norm =

34000 with $\alpha = 0.29$. In Fig. 13A, the original *1024×1024* Magnetic Tape sample has TV norm = 14000 with $\alpha = 0.35$. The APEX-processed Fig. 13B has TV norm = 39000 with $\alpha = 0.26$. These very substantial sharpness increases are typical of those obtained in numerous other test images.

Measuring Image Smoothness

Most commonly occurring images $f(x,y)$ are not differentiable functions of the variables x and y. Rather, these images display edges, localized sharp features, and other fine-scale details or *texture*. Many digital image-processing tasks require prior specification of the correct mathematical function space in which the true image lies. If an image is incorrectly postulated to be too smooth, the processing algorithm may produce an overly smoothed version of the true image in which critical information has been lost.

During the last 10 years, a very considerable amount of image analysis research has been based on the assumption that most images belong to the space of functions of bounded variation. However, it has been subsequently discovered that such so-called total variation (TV) image processing sometimes results in unacceptable loss of fine-scale information. This phenomenon is now known as the *staircase effect*. In papers published in 2001, French researchers Gousseau, Morel, and Meyer, showed that most natural images are, in fact, not of bounded variation, and that TV image-processing techniques must inevitably smooth out texture.

Correct characterization of the lack of smoothness of images is a fundamental problem in image processing. It turns out that so-called *Lipschitz spaces* are the appropriate framework for accommodating non-smooth images. The L^P Lipschitz exponent α for the given image, where $0 < \alpha < 1$, measures the fine-scale content of that image, provided the image is relatively noise free. Heavily textured imagery has low values for α, while large values of α indicate that the image is relatively smooth. Estimating an image's Lipschitz exponent is a delicate problem. We have developed a new, computationally efficient, method for estimating α. It merely requires blurring the image by convolution with a specific singular integral kernel, and evaluating the discrete L^P norm of the difference between the blurred and original images. The rate at which this L^P norm tends to zero, as the kernel approaches the Dirac δ-function, is directly related to the Lipschitz exponent α. Since the required convolutions can be accomplished by FFTs, very minimal computational effort is thus needed to implement the resulting procedure. In addition, this approach has the advantage of allowing consideration of substantially wider Lipschitz spaces than is mathematically possible with existing procedures, thereby encompassing a much wider class of images.

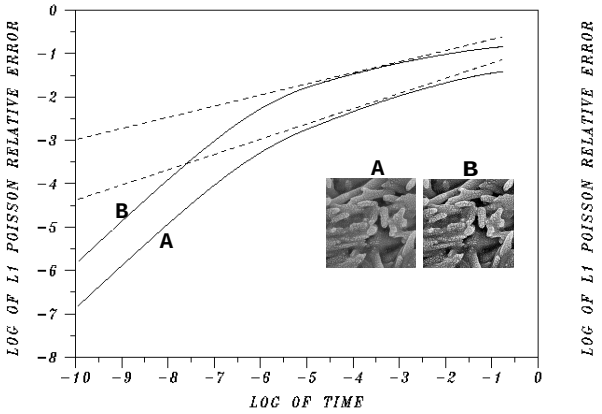

Figure 14. *The Lipschitz exponent α in an image can be obtained by plotting* $\log(\|U^t f - f\| / \|f\|)$ *versus* $\log t$ *as t tends to zero, where U^t is the Poisson kernel with width t. The slope of dashed line which asymptotes the solid trace for* $-6 < \log t < 0$ *is the Lipschitz exponent. Here, the original SEM image A has $\alpha = 0.35$. The APEX processed image B reveals more fine scale structure and has lower $\alpha = 0.26$. (See Fig. 13 for larger images of A and B.)*

Significant potential applications of this technology include the routine monitoring of image sharpness and imaging performance in electro-optical imaging systems, the performance evaluation of image reconstruction software, the detection of possibly abnormal fine-scale features in some medical imaging applications, and the monitoring of surface finish in industrial applications. In addition, specifying the correct Lipschitz space wherein an image lies can be used to solve the blind image deconvolution problem in a way that preserves texture, i.e., fine detail, in a recovered image. We recently have developed a new method, the Poisson Singular Integral (PSI) method, which yields an excellent approximation to optimal image filtering for a very wide class of images.

References

A. Carasso, Singular Integrals, Image Smoothness, and the Recovery of Texture in Image Deblurring, *SIAM Journal on Applied Mathematics*, **64** (5), (2004), pp. 1749-1774.

Participants

A. Carasso (MCSD); A. Vladár (MEL)

Creating Visual Models of Nanoworlds

The place where quantum and macro effects meet, the nanoworld is full of the unexpected. In such an environment, visual models of laboratory and computational experiments can be critical to comprehension. We are developing an immersive visual modeling system that enables scientists to easily view and interact with their data in multiple ways in real time.

Howard Hung, Steve Satterfield, James Sims, Adele Peskin, John Kelso, John Hagedorn, Terrence Griffin, and Judith Terrill

---

Computational and laboratory experiments are generating increasing amounts of scientific data. Often, the complexity of the data makes it difficult to devise *a priori* methods for its analysis, or the data is from new landscapes, such as the nanoworld, where we have little experience. Moreover, there may be ancillary data, from databases for example, that would be helpful to have available. We are developing visual analysis capabilities in an immersive environment that allow scientists to interact with data objects in a three-dimensional landscape rather than simply viewing pictures of them. Fully immersive computer graphics include one or more large rear projection screens to encompass peripheral vision, stereoscopic display for increased depth perception, as well as head tracking for realistic perspective based on the direction the user is viewing. With visual exploration, scientists can easily perceive complex relationships in their data, quickly ascertaining whether the results match expectations. Real time interaction adds to the potential for speeding the process of insight.

In a fast moving field like nanotechnology, it is important to be able to create and interact with new visual models quickly. Our visual environment is built for generality, flexibility and speed. Rather than a single monolithic program, it is a collection of tools designed to work together to create, display, and interact with visual models. We have created three main categories of tools: infrastructure software, representation software, and scene interaction software. We join programs together using Unix pipes and filters for creation and transformation. We construct Dynamically Shared Objects (DSO's) for functionality, and scenegraph objects for ease of placement.

DIVERSE

Our visual environment centers around a core infrastructure program called DIVERSE (Device Independent Visualization Environment – Reconfigurable, Scalable, Extensible). DIVERSE, which was developed by Virginia Tech with support and technical contributions by NIST, is an interface that facilitates the development of immersive computer graphics programs for use on a wide variety of graphics displays. DIVERSE provides a toolkit to load in previously compiled objects, called DSOs. DSOs can be used to describe the graphics display, input devices, navigation techniques or interaction styles for a given visualization. Using collections of DSOs, applications can be reconfigured without recompiling. The same DSO that defines how a wand controls an object can be used both in an immersive environment and on a laptop computer. With the ability to import the output of conventional visualization programs, and access to the increasing capabilities of commodity graphics cards, we have a very rich environment in which to express and communicate visual models.

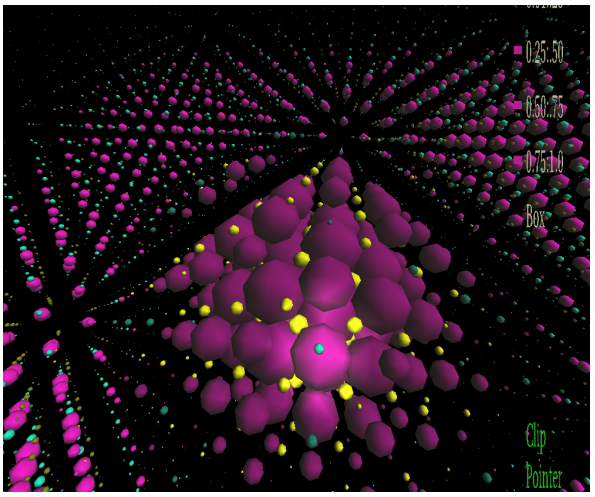

Figure 15. *Image of a quantum dot created from the output of a computer simulation of the optical properties of nanostructures. The spheres represent s-orbitals.*

Representation

Representation refers to the process of transforming raw data into a visual geometric format that can be viewed and manipulated. The key to quickly and easily visualizing scientific data in an immersive environment is the capability provided by the set of tools used to convert raw data to immersive data. We have created a variety of tools to make glyphs, color them, place them, and render them to a desired level of transparency and other properties. While our system is based on our own internal representation, we take

advantage of representations computed by other packages by providing tools to take the output of common packages and convert them into the format that places them in our environment.

Interaction

For efficient scientific exploration, it is important to have user interaction that is both easy to set up and easy to adapt to differing needs. Individual DSOs that add new functionality to our visualization software system are loaded at run-time, and their behavior is cumulative. Together, a set allows for a huge range of combinations data visualization modes. We have developed a wide range of DSOs that allow a user to interact with the objects to be viewed. This includes functionality to move objects around, select individual objects or sets, assign functionality to the selections, interact with outside software, bring data into the system, send data out of the system, and load or unload objects during visualization. They can also interactively select the level of detail in a scene. Individual DSOs can add simple capability to a scenegraph, such as adding a particular light source or an object to represent a pointer for the user to select objects. DSOs can also add tasks to the objects of a scenegraph. When selected, an object can be given the task of turning on or off another object or itself, or executing a command to interact with another DSO or an external program. Objects can change visibility, initiate new simulations, define clipping planes, and so on. DSOs provide a variety of ways to navigate through the environment, including changing the scale at will. They allow viewing of individual scenes as well as time sequences (i.e., movies). The output of an interaction can be saved as a simple image, as a movie, or the interaction itself can be saved as an experience that can be replayed at a later time.

Applications to the Nanoworld

We have successfully applied our visual analysis techniques and tools to the study of a variety of nanoscale phenomena. Among these are visual analyses of

- s-orbitals for the simulation of electronic and optical properties of complex nanostructures such as semiconductor nanocrystals and quantum dots (see Fig. 15),
- electric, magnetic, and energy field vectors from the simulation of optical scattering by metamagnetic materials (see Fig. 16),
- intermediate voltage electron microscope measurement approaches to attain three-dimensional chemical images at nanoscale-resolution, and
- dynamics of molecular interactions leading to the formation of smart gels.

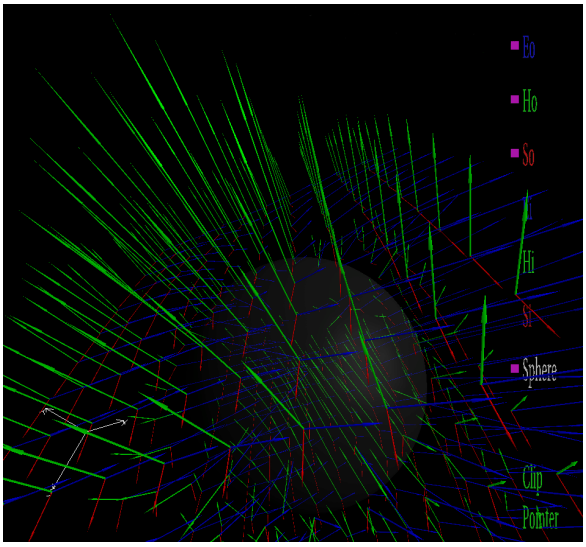

Figure 16. *Image of the electric, magnetic, and energy field vectors in a single time step in the simulation of resonant optical scattering by metamagnetic materials.*

References

J.E. Devaney, S.G. Satterfield, J.G. Hagedorn, J.T. Kelso, A.P. Peskin, W.L. George, T.J. Griffin, H.K. Hung, R.D. Kriz, Science at the Speed of Thought, in *Ambient Intelligence for Scientific Discovery*, (Y. Cai, ed.), Lecture Notes in Artificial Intelligence **3345**, February 2005.

J. Sims, G. Bryant, H. Hung, Intrinsic Surface States in Semiconductor Nanocrystals: HgS Quantum Dots, American Physical Society, Los Angeles, CA, March 21-25, 2005.

J.P. Dunkers, J.G. Hagedorn, A. Peskin, J.T. Kelso, J.E. Terrill, and L. Henderson, Interactive, Quantitative Analysis of Scaffold Structure Using Immersive Visualization, *BIO2006: Summer Bioengineering Conference*.

Participants

J. Terrill, S. Satterfield, J. Kelso, J. Hagedorn, A. Peskin, H. Hung, T. Griffin, A. Dienstfrey (MCSD); G. Bryant, Z. Levine (PL); C. Gonzales (CSTL), J. Dunkers, L. Henderson (MSEL); R. Kriz (Virginia Tech)

`http://math.nist.gov/mcsd/savg/vis/`

Measurement Science in the Virtual World

We are developing tools for selection and measurement from within the immersive visualization environment. These visualizations and their accompanying analyses then yield quantitative results that extend the qualitative knowledge that is the typical product of visualization.

John Hagedorn, Adele Peskin, John Kelso, Steve Satterfield, and Judith Terrill

Laboratory experiments, computational experiments, measurement, analysis, and visualization are typically separate activities. This slows down the rate at which knowledge is gained. Combining them, where feasible, yields greater efficiency, and the resulting synergy can deepen understanding. We are working on combining measurement, real time immersive visualization, and analysis. Our immersive visualization environment is the main tool which enables this.

It is not always possible to perform desired measurements on laboratory data during the course of a physical experiment. For example, a reconstruction phase may be needed to get the data into a form where it can be measured. An example is tomographic reconstruction. The same is true for computational experiments. While measurements are typically gathered during runs, there are also many reasons why measurements may need to be taken during the analysis stage. Measurement from within the immersive environment provides unique capabilities in this regard. We are working to use these capabilities to enable such virtual measurements to be taken both during and after laboratory and computational experiments.

To bring measurement science into the virtual world we need the objects in the environment to be drawn and positioned with higher precision than is typical in applications of immersive virtual reality. To do this we first need the virtual space to be calibrated. Then we need to be able to select and measure properties of objects. Finally, we need to be able to operate on those measurements and relate them to both the immersive environment and the real world. We describe our approach to each of these in turn.

Calibration

The use of motion tracking devices is essential in immersive visualization systems. The location and orientation of the user's eyes must be continually tracked so that the system can render images of the virtual world, in stereo, in real-time. In addition, motion tracking is also commonly used to track hand-held devices (e.g., a wand) that operate as tools in the virtual world created within the immersive display.

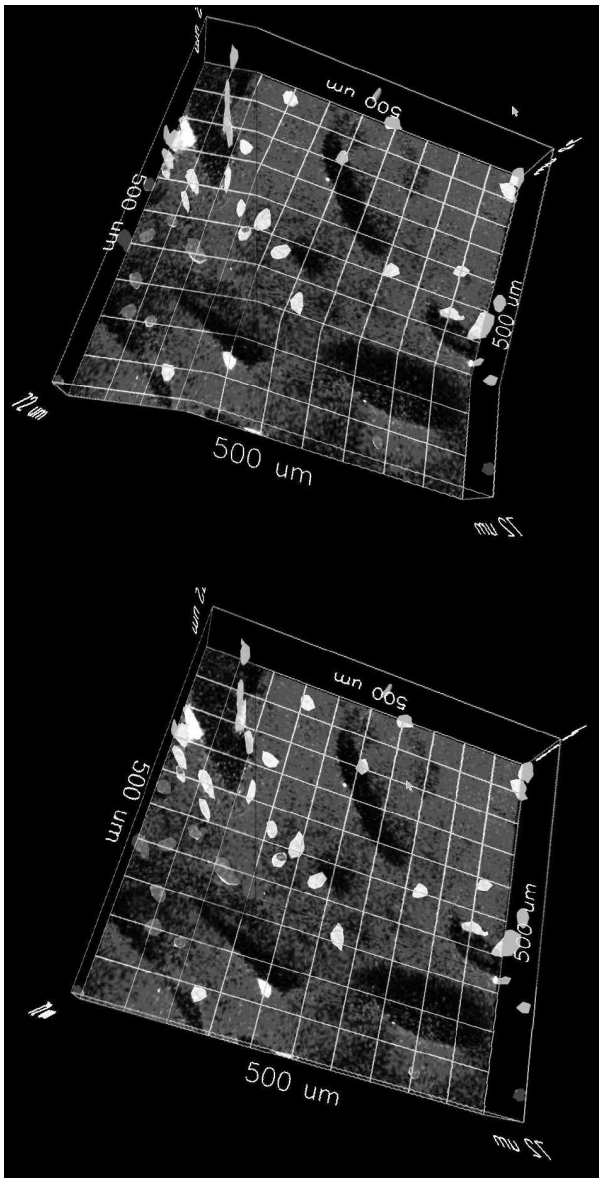

Figure 17. *The top image shows visual artifacts due to tracker error. The bottom image shows the same scene with tracker data corrected.*

Environmental factors such as the presence of metallic objects and electromagnetic fields generated by video monitors contribute substantially to errors in measurements reported by electromagnetic tracking devices. We have developed a method for calibrating and correcting location and orientation errors from electromagnetic motion tracking devices. This includes a new algorithm for interpolating rotation corrections at

scattered data points. This method, unlike previous methods, is rooted in the geometry of the space of rotations. Results have shown large improvements in the precision of both location and orientation measurements. The methods impose minimal computational burden.

Selection, Measurement, and Analysis

We have developed a capability for linear measurements in the immersive environment. Our objective is to build a software system within the environment that integrates the following tasks: (1) linear measurement, (2) analysis of the collected measurements, (3) display of the results, and (4) interactions with the data and analyses that will enable grouping of results. The goal of these tasks is to achieve greater understanding by generating quantitative results.

There are two main components to our implementation. The first is software that allows the user to manually make a series of linear measurements in the immersive environment. The second is a standardized 2D user interface displaying the measurement statistics and distribution in tabular and histogram form.

Our main objective in designing the user interface was to make the 3D measurement task direct and natural. The user makes a linear measurement simply by moving the hand-held wand to a point in the 3D virtual space, pressing a button on the device, then moving to a second point and pressing the button again. Visual feedback is given at each step of the process and the user is able to adjust each end point simply by grabbing it with the wand and repositioning it. The process is fast, simple, and intuitive. A user interface enables one to bring the measurements into an analysis and to display the results. In Fig. 18, the user has made measurements (shown in blue). The measurements have been brought out of the environment and a mean and standard deviation computed. They are also shown as a histogram in the upper right. Portions of the histogram can be selected and highlighted back in the immersive environment for a better understanding.

We have found that immersive visualization enables both qualitative and quantitative understanding of 3D structure that was not otherwise possible. These types of measurements made in the virtual environment would be very difficult to make with typical desktop visualization techniques. We are planning a variety of extensions to our selection and measurement capability for the future.

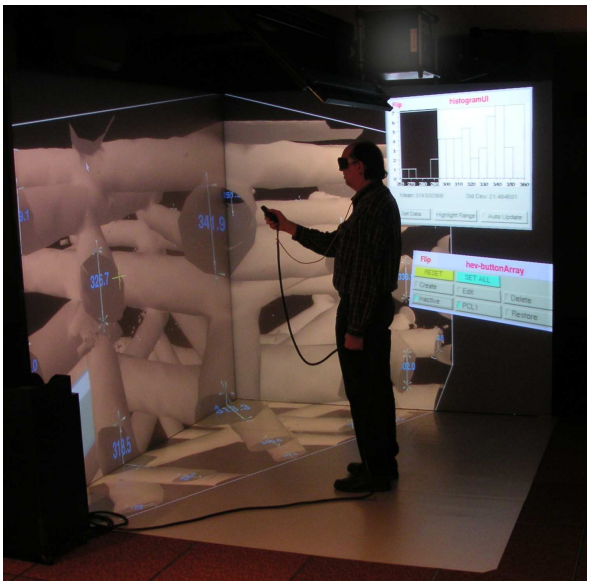

Figure 18. *A user in the NIST immersive visualization environment making measurements on a scaffold for tissue engineering. The measurements are shown in blue. The histogram of results is shown in the upper right.*

Current Applications

We are currently working on a variety of applications spanning several disciplines.

- Reference scaffold for Tissue Engineered materials
- Correlate microstructure with global properties of hydrating cement
- Particle identity and distribution at the nanoscale
- Stress and strain at the nanoscale

References

J. Hagedorn, S. Satterfield, J. Kelso, W. Austin, J. Terrill, and A. Peskin, Correction of Location and Orientation Errors in Electromagnetic Motion Tracking, *Presence*, submitted.

J. Hagedorn, J. Dunkers, A. Peskin, J. Kelso, L. Henderson, and J. Terrill, Quantitative, Interactive Measurement of Tissue Engineering Scaffold Structure in an Immersive Visualization Environment, *2006 IEEE International Symposium on Biomedical Imaging*, submitted.

Participants

J. Hagedorn, A. Peskin, J. Kelso, S. Satterfield, J. Terrill (MCSD); J. Dunkers, L. Henderson (MSEL)

Identifying Objects in LADAR Scanning Data

Laser scanning technology has developed into a major tool for geographic and geometric data acquisition. We are investigating the feasibility of using such technology for monitoring construction sites. A core problem in the analysis of such data is object recognition. Recently we demonstrated a technique for identifying particular objects, such as I-beams in noisy LADAR scanning data.

David Gilsinn

During the past decade, laser-scanning technology has developed into a major vehicle for wide-ranging applications such as cartography, bathymetry, urban planning, object detection, and dredge volume determination. One advantage of such scanning data is that it can provide information on spatial relationships that ordinary photography cannot. The NIST Building and Fire Research Lab (BFRL) is investigating the use of such technology to monitor progress of work on construction sites. Here laser scans taken from several vantage points are used to construct a surface model representing a particular scene. A library of 3D representations of construction site objects, obtained from CAD data sets, would also be used. The objects would be loaded into an associated simulation system that tracks both equipment and resources based on real-time data from the construction site obtained from laser scans.

Pick-and-place control of construction site objects is a major application. With automation and robotics entering on construction site scene, vision systems, such as LADAR (laser direction and ranging), will be incorporated for real time object identification, based on 3D library templates. Once objects, such as I-beams, are located, robotic crane grippers can be manipulated to acquire the I-beam. We are developing and testing algorithms for this particular task.

LADAR scans of I-beams of multiple lengths and angular poses relative to the scanning LADAR have been generated (see Fig. 19). A database of design specifications for potential I-beam candidates has been created. The LADAR scans generate a large number of points, ranging in the millions, which can be acquired seconds. These scans usually contain a large number of noisy data values arising from ground hits to phantom pixels caused by beam splitting at sharp edges.

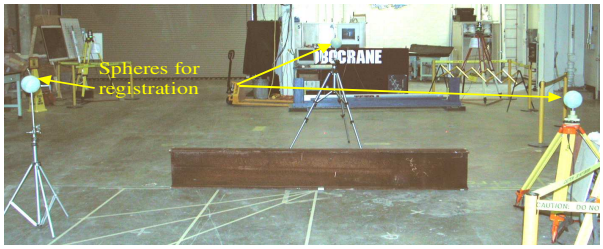

Figure 19. *Experimental configuration for scanning an I-beam by a LADAR placed at the camera location. The white lines on the floor are alignment marks for setting the I-beam at different angles relative to the LADAR. The spheres on the tripods are located for coordinate registration with a world coordinate system.*

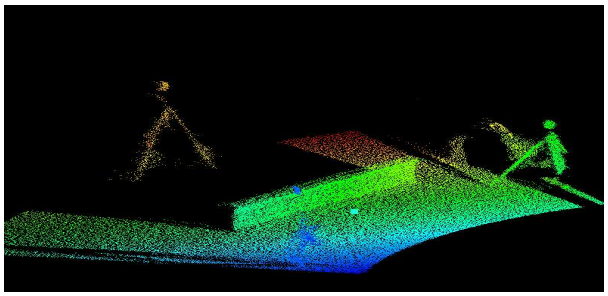

Figure 20. *The results of a scan of the I-beam in Fig. 19. The flat areas on either side of what appears to be the I-beam are floor hits. Notice the results of hits on the three tripods, one of which is seen as somewhat of a shadow in the lower middle of the scene. Also note the occlusion shadowing.*

Fig. 20 shows a typical LADAR scan. Whereas the photograph in Fig. 19 provides a clear image it does not provide coordinate information. Although each of the points in the figure are associated with an (*x, y, z*) coordinate relative to the LADAR, the identification of the nature of the objects scanned is quite difficult. The challenge is to use the database of design specifications, in this case of I-beams, to locate an object in the scanned image and to report its center-of-mass location and angular pose relative to the LADAR.

Three main tasks must be performed: segmenting the data points into groups corresponding to likely objects, identifying the segmented objects with the highest likelihood of being the object sought, and registering the object relative to a world coordinate system so that the LADAR scan can be related to a global coordinate system in which other objects are potentially located.

Segmentation Algorithms

Two algorithms have been developed. The first uses binning of the scanned points to reduce data volume. The bins are then examined to identify those that are likely to be phantom points or floor hits. Phantom points are coordinate points generated by the LADAR due to the internal process of averaging the returned beam signal. Due to the finite size of a LADAR beam,

a targeted point at the edge of an object can lead to a partial signal returns from both the object and a more distant one. The entire signal is averaged and a coordinate point returned somewhere between the two.

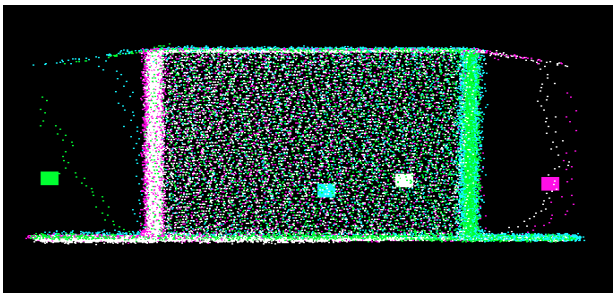

Figure 21. *Anomalies from scanning a box from different directions. Phantom points off of the top edges are seen as two lines beginning at the edges. Phantom points along the sides of the box are due to the averaging of LADAR hits along the sides of the box.*

In Fig. 21 a line of phantom points is seen emanating from the top edges of the test box. Phantom points due to hits along the edges are also visible. Bins that include phantom points and floor hits can be eliminated since they are likely to be sparsely populated. Object identification occurs when the bins are grouped into potential objects and bounding boxes are placed around them. These boxes are compared with bounding boxes defining I-beams in the database and the best fit is reported, along with the center-of-mass and pose of the bounding box. An example of the bounding box placed around the I-beam is shown in Fig. 22.

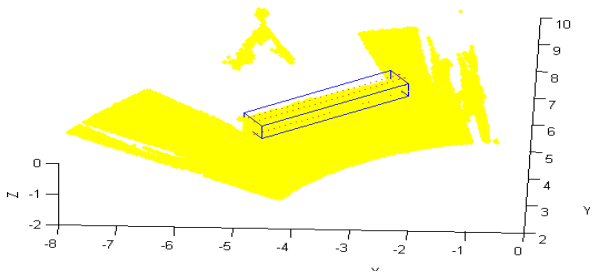

Figure 22. *Bounding box around the I-beam scanned in Fig 20.*

The second algorithm uses Triangulated Irregular Networks (TINs) to mesh the data. The density of the triangulated points is then visually examined to identify those triangle groups most likely to form objects. A TIN of the data in Fig. 20 is illustrated in Fig. 23. The bounding box procedure for object identification is then applied as before. Further work is needed with this algorithm to eliminate manual density inspection.

Evaluation

In order to determine the accuracy of the predicted pose of the I-beams, measurements at the four upper corner points of the I-beams were made using a laser-based Site Metrology System (SMS) developed by BFRL. These reference points were used to measure the performance of the algorithms in their ability to locate the points. The binning algorithm performed well in identifying length of the I-beam, the location of the center of the I-beam, and the angle relative to the LADAR scan direction until the I-beam's major axis was placed at 30 degrees and 0 degrees relative to the scan direction. Thus as the major axis of the I-beam became more aligned with the scan direction of the LADAR the algorithm had a harder time identifying the I-beam. The TIN algorithm performed better than the binning algorithm as the I-beam axis became aligned with the scan beam direction. However, as in the binning algorithm the TIN algorithm could not identify the I-beam when the major axis was aligned with the scan beam.

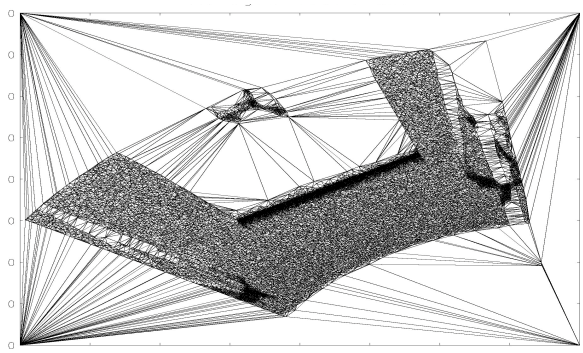

Figure 23. *TIN image of the data in Fig 20. The density of triangles can be used to locate the I-beam and reference sphere tripods.*

Future work will involve extending the bounding box concept to other shapes such as polygons, cylinders, etc. in order to develop a general comparison algorithm that covers most objects located at construction sites, such as vehicles, gas cylinders, etc.

References

D.E. Gilsinn, G.S. Cheok, C. Witzgall, and A. Lytle, Construction Object Identification from LADAR Scans: An Experimental Study Using I-beams, NIST Interagency Report, 2005.

D.E. Gilsinn, G.S. Cheok, and D. O'Leary, Reconstructing Images of Bar Codes for Constructions Site Object Recognition, *Automation in Construction*, **13** (2004) 21-35.

Participants

D. Gilsinn, C. Witzgall, D. O'Leary (MCSD); G. Cheok, A. Lytle, W. Stone (BFRL)